



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/717,473

DATE: 08/05/2004

TIME: 08:55:45

Input Set : N:\CrF3\RULE60\10717473.raw  
 Output Set: N:\CRF4\08052004\J717473.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

- 5 (i) APPLICANT: WEI, Ying-Fei
- 7 (ii) TITLE OF INVENTION: Transforming Growth Factor
- 8 Alpha HIII
- 9 (iii) NUMBER OF SEQUENCES: 10
- 10 (iv) CORRESPONDENCE ADDRESS:
  - 11 (A) ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
  - 12 Cecchi, Stewart & Olstein
  - 13 (B) STREET: 6 Becker Farm Road
  - 14 (C) CITY: Roseland
  - 15 (D) STATE: NJ
  - 16 (E) COUNTRY: USA
  - 17 (F) ZIP: 07068-1739
- 19 (v) COMPUTER READABLE FORM:
  - 20 (A) MEDIUM TYPE: Floppy disk
  - 21 (B) COMPUTER: IBM PC compatible
  - 22 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - 23 (D) SOFTWARE: WordPerfect 5.1, Dos Text File

## 25 (vi) CURRENT APPLICATION DATA:

- C--> 26 (A) APPLICATION NUMBER: US/10/717,473
- C--> 27 (B) FILING DATE: 21-Oct-2003
- W--> 33 (C) CLASSIFICATION: 435

## 35 (vii) PRIOR APPLICATION DATA:

- 31 (A) APPLICATION NUMBER: US/08/778,545
- 32 (B) FILING DATE: January 3, 1997
- 36 (A) APPLICATION NUMBER: 60/011,136
- 37 (B) FILING DATE: January 4, 1996

## 40 (viii) ATTORNEY/AGENT INFORMATION:

- 41 (A) NAME: MULLINS, J.G.
- 42 (B) REGISTRATION NUMBER: 33,073
- 43 (C) REFERENCE/DOCKET NUMBER: 325800-541 (PF 220)

## 45 (ix) TELECOMMUNICATION INFORMATION:

- 46 (A) TELEPHONE: 201-994-1700
- 47 (B) TELEFAX: 201-994-1744

## 49 (2) INFORMATION FOR SEQ ID NO: 1:

## 51 (i) SEQUENCE CHARACTERISTICS:

- 52 (A) LENGTH: 923 base pairs
- 53 (B) TYPE: nucleic acid
- 54 (C) STRANDEDNESS: double
- 55 (D) TOPOLOGY: linear

## 57 (ii) MOLECULE TYPE: DNA (genomic)

## 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:



RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/717,473

DATE: 08/05/2004  
TIME: 08:55:45

Input Set : N:\CrF3\RULE60\10717473.raw  
Output Set: N:\CRF4\08052004\J717473.raw

62	GAAA	ATG	GCG	CCT	CAC	GGC	CCG	GGT	AGT	CTT	ACG	ACC	CTG	GTG	CCC	TGG	49
63	Met	Ala	Pro	His	Gly	Pro	Gly	Ser	Leu	Thr	Thr	Leu	Val	Pro	Trp		
64	-25				-20				-15								
66	GCT	GCC	GCC	CTG	CTC	CTC	GCT	CTG	GGC	GTG	GAA	AGG	GCT	CTG	GCG	CTA	97
67	Ala	Ala	Ala	Leu	Leu	Leu	Ala	Leu	Gly	Val	Glu	Arg	Ala	Leu	Ala	Leu	
W--> 68	-10			-5					1				5				
70	CCC	GAG	ATA	TGC	ACC	CAA	TGT	CCA	GGG	AGC	GTG	CAA	AAT	TTG	TCA	AAA	145
71	Pro	Glu	Ile	Cys	Thr	Gln	Cys	Pro	Gly	Ser	Val	Gln	Asn	Leu	Ser	Lys	
72		10				15				20							
74	GTG	GCC	TTT	TAT	TGT	AAA	ACG	ACA	CGA	GAG	CTA	ATG	CTG	CAT	GCC	CGT	193
75	Val	Ala	Phe	Tyr	Cys	Lys	Thr	Thr	Arg	Glu	Leu	Met	Leu	His	Ala	Arg	
76		25				30			35								
78	TGC	TGC	CTG	AAT	CAG	AAG	GGC	ACC	ATC	TTG	GGG	CTG	GAT	CTC	CAG	AAC	241
79	Cys	Cys	Leu	Asn	Gln	Lys	Gly	Thr	Ile	Leu	Gly	Leu	Asp	Leu	Gln	Asn	
80		40				45			50								
82	TGT	TCT	CTG	GAG	GAC	CCT	GGT	CCA	AAC	TTT	CAT	CAG	GCA	CAT	ACC	ACT	289
83	Cys	Ser	Leu	Glu	Asp	Pro	Gly	Pro	Asn	Phe	His	Gln	Ala	His	Thr	Thr	
84		55			60			65		70							
86	GTC	ATC	ATA	GAC	CTG	CAA	GCA	AAC	CCC	CTC	AAA	GGT	GAC	TTG	GCC	AAC	337
87	Val	Ile	Ile	Asp	Leu	Gln	Ala	Asn	Pro	Leu	Lys	Gly	Asp	Leu	Ala	Asn	
88			75				80			85							
90	ACC	TTC	CGT	GGC	TTT	ACT	CAG	CTC	CAG	ACT	CTG	ATA	CTG	CCA	CAA	CAT	385
91	Thr	Phe	Arg	Gly	Phe	Thr	Gln	Leu	Gln	Thr	Leu	Ile	Leu	Pro	Gln	His	
92		90				95				100							
94	GTC	AAC	TGT	CCT	GGG	GGG	ATT	AAT	GCC	TGG	AAT	ACT	ATC	ACC	TCT	TAT	433
95	Val	Asn	Cys	Pro	Gly	Gly	Ile	Asn	Ala	Trp	Asn	Thr	Ile	Thr	Ser	Tyr	
96		105			110			115									
98	ATA	GAC	AAC	CAA	ATC	TGT	CAA	GGG	CAA	AAG	AAC	CTT	TGC	AAT	AAC	ACT	481
99	Ile	Asp	Asn	Gln	Ile	Cys	Gln	Gly	Gln	Lys	Asn	Leu	Cys	Asn	Asn	Thr	
100		120				125			130								
102	GGG	GAC	CCA	GAA	ATG	TGT	CCT	GAG	AAT	GGA	TCT	TGT	GTA	CCT	GAT	GGT	529
103	Gly	Asp	Pro	Glu	Met	Cys	Pro	Glu	Asn	Gly	Ser	Cys	Val	Pro	Asp	Gly	
104		135			140			145			150						
106	CCA	GGT	CTT	TTG	CAG	TGT	TGT	GCT	GAT	GGT	TTC	CAT	GGA	TAC	AAG	577	
107	Pro	Gly	Leu	Leu	Gln	Cys	Val	Cys	Ala	Asp	Gly	Phe	His	Gly	Tyr	Lys	
108			155				160			165							
110	TGT	ATG	CGC	CAG	GGC	TCG	TTC	TCA	CTG	CTT	ATG	TTC	TTC	GGG	ATT	CTG	625
111	Cys	Met	Arg	Gln	Gly	Ser	Phe	Ser	Leu	Leu	Met	Phe	Phe	Gly	Ile	Leu	
112		170				175			180								
114	GGA	GCC	ACC	ACT	CTA	TCC	GTC	TCC	ATT	CTG	CTT	TGG	GCG	ACC	CAG	CGC	673
115	Gly	Ala	Thr	Thr	Leu	Ser	Val	Ser	Ile	Leu	Leu	Trp	Ala	Thr	Gln	Arg	
116		185				190			195								
118	CGA	AAA	GCC	AAG	ACT	TCA	TGA	ACT	ACAT	AGGT	CTT	ACC	ATT	GAC	CTAA	721	
119	Arg	Lys	Ala	Lys	Thr	Ser											
120		200															
122	GAT	CAAT	CTG	AACT	ATCTTA	GCCCAGTCAG	GGAGCTCTGC	TTC	CCTAGAAA	GGCAT	CTTTTC						781
124	GCC	AGT	GGAT	TCG	CCTCAAG	GTTGAGGCCG	CCATTGGAAG	ATG	AAA	ATT	GCAC	TC	CTT				841
126	GGT	GTA	GACA	AAT	ACCAGTT	CCCATTGGTG	TTGTTGCCTA	TAATAAACAC	TTTTT	CTTT							901
128	TTT	AAAAAAA	AAAAAAA	AA													923

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/717,473

DATE: 08/05/2004  
TIME: 08:55:45

Input Set : N:\Crf3\RULE60\10717473.raw  
Output Set: N:\CRF4\08052004\J717473.raw

130 (2) INFORMATION FOR SEQ ID NO: 2:  
132 (i) SEQUENCE CHARACTERISTICS:  
133 (A) LENGTH: 229 amino acids  
134 (B) TYPE: amino acid  
135 (D) TOPOLOGY: linear  
137 (ii) MOLECULE TYPE: protein  
139 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
141 Met Ala Pro His Gly Pro Gly Ser Leu Thr Thr Leu Val Pro Trp Ala  
142 -25 -20 -15 -10  
144 Ala Ala Leu Leu Ala Leu Gly Val Glu Arg Ala Leu Ala Leu Pro  
145 -5 1 5  
147 Glu Ile Cys Thr Gln Cys Pro Gly Ser Val Gln Asn Leu Ser Lys Val  
148 10 15 20  
150 Ala Phe Tyr Cys Lys Thr Thr Arg Glu Leu Met Leu His Ala Arg Cys  
151 25 30 35  
153 Cys Leu Asn Gln Lys Gly Thr Ile Leu Gly Leu Asp Leu Gln Asn Cys  
154 40 45 50 55  
156 Ser Leu Glu Asp Pro Gly Pro Asn Phe His Gln Ala His Thr Thr Val  
157 60 65 70  
159 Ile Ile Asp Leu Gln Ala Asn Pro Leu Lys Gly Asp Leu Ala Asn Thr  
160 75 80 85  
162 Phe Arg Gly Phe Thr Gln Leu Gln Thr Leu Ile Leu Pro Gln His Val  
163 90 95 100  
165 Asn Cys Pro Gly Gly Ile Asn Ala Trp Asn Thr Ile Thr Ser Tyr Ile  
166 105 110 115  
168 Asp Asn Gln Ile Cys Gln Gly Gln Lys Asn Leu Cys Asn Asn Thr Gly  
169 120 125 130 135  
171 Asp Pro Glu Met Cys Pro Glu Asn Gly Ser Cys Val Pro Asp Gly Pro  
172 140 145 150  
174 Gly Leu Leu Gln Cys Val Cys Ala Asp Gly Phe His Gly Tyr Lys Cys  
175 155 160 165  
177 Met Arg Gln Gly Ser Phe Ser Leu Leu Met Phe Phe Gly Ile Leu Gly  
178 170 175 180  
180 Ala Thr Thr Leu Ser Val Ser Ile Leu Leu Trp Ala Thr Gln Arg Arg  
181 185 190 195  
183 Lys Ala Lys Thr Ser  
184 200  
187 (2) INFORMATION FOR SEQ ID NO: 3:  
189 (i) SEQUENCE CHARACTERISTICS:  
190 (A) LENGTH: 27 base pairs  
191 (B) TYPE: nucleic acid  
192 (C) STRANDEDNESS: single  
193 (D) TOPOLOGY: linear  
195 (ii) MOLECULE TYPE: DNA (genomic)  
197 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
199 CGCGGATCCG GGCAAAAGAA CCTTTGC 27  
202 (2) INFORMATION FOR SEQ ID NO: 4:  
204 (i) SEQUENCE CHARACTERISTICS:  
205 (A) LENGTH: 30 base pairs

RAW SEQUENCE LISTING DATE: 08/05/2004  
 PATENT APPLICATION: US/10/717,473 TIME: 08:55:45

Input Set : N:\Crf3\RULE60\10717473.raw  
 Output Set: N:\CRF4\08052004\J717473.raw

206 (B) TYPE: nucleic acid  
 207 (C) STRANDEDNESS: single  
 208 (D) TOPOLOGY: linear  
 210 (ii) MOLECULE TYPE: DNA (genomic)  
 212 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 214 GCGTCTAGAC TAAAGCAGTG AGAACGAGCC 30  
 217 (2) INFORMATION FOR SEQ ID NO: 5:  
 219 (i) SEQUENCE CHARACTERISTICS:  
 220 (A) LENGTH: 34 base pairs  
 221 (B) TYPE: nucleic acid  
 222 (C) STRANDEDNESS: single  
 223 (D) TOPOLOGY: linear  
 225 (ii) MOLECULE TYPE: DNA (genomic)  
 227 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 229 CGCGGATCCG TCCATCATGG CGCCTCACGG CCCG 34  
 232 (2) INFORMATION FOR SEQ ID NO: 6:  
 234 (i) SEQUENCE CHARACTERISTICS:  
 235 (A) LENGTH: 33 base pairs  
 236 (B) TYPE: nucleic acid  
 237 (C) STRANDEDNESS: single  
 238 (D) TOPOLOGY: linear  
 240 (ii) MOLECULE TYPE: DNA (genomic)  
 242 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 244 GCGTCTAGAC TACATAAGCA GTGAGAACGA GCC 33  
 248 (2) INFORMATION FOR SEQ ID NO: 7:  
 250 (i) SEQUENCE CHARACTERISTICS:  
 251 (A) LENGTH: 28 base pairs  
 252 (B) TYPE: nucleic acid  
 253 (C) STRANDEDNESS: single  
 254 (D) TOPOLOGY: linear  
 256 (ii) MOLECULE TYPE: DNA (genomic)  
 258 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
 260 CGCGGATCCC GGGCAAAAGA ACCTTTGC 28  
 263 (2) INFORMATION FOR SEQ ID NO: 8:  
 265 (i) SEQUENCE CHARACTERISTICS:  
 266 (A) LENGTH: 33 base pairs  
 267 (B) TYPE: nucleic acid  
 268 (C) STRANDEDNESS: single  
 269 (D) TOPOLOGY: linear  
 271 (ii) MOLECULE TYPE: DNA (genomic)  
 273 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
 275 GCGTCTAGAC TACATAAGCA GTGAGAACGA GCC 33  
 278 (2) INFORMATION FOR SEQ ID NO: 9:  
 280 (i) SEQUENCE CHARACTERISTICS:  
 281 (A) LENGTH: 30 base pairs  
 282 (B) TYPE: nucleic acid  
 283 (C) STRANDEDNESS: single  
 284 (D) TOPOLOGY: linear  
 286 (ii) MOLECULE TYPE: DNA (genomic)

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/717,473

DATE: 08/05/2004  
TIME: 08:55:45

Input Set : N:\Crf3\RULE60\10717473.raw  
Output Set: N:\CRF4\08052004\J717473.raw

288 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
290 GCGCTCAGAC ATAAGCAGTG AGAACGAGCC 30  
293 (2) INFORMATION FOR SEQ ID NO: 10:  
295 (i) SEQUENCE CHARACTERISTICS:  
296 (A) LENGTH: 52 amino acids  
297 (B) TYPE: amino acid  
298 (C) STRANDEDNESS: single  
299 (D) TOPOLOGY: linear  
301 (ii) MOLECULE TYPE: protein  
303 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
305 Gln Gly Leu Asn Lys Cys Asn Asn Thr Gly Asp Pro Glu Met Cys Pro  
306 1 5 10 15  
308 Glu Asn Gly Ser Cys Val Pro Asp Gly Pro Gly Leu Leu Gln Cys Val  
309 20 25 30  
311 Cys Ala Asp Gly Phe His Gly Tyr Lys Cys Met Arg Gln Gly Ser Phe  
312 35 40 45  
314 Ser Leu Leu Met  
315 50

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/717,473

DATE: 08/05/2004

TIME: 08:55:46

Input Set : N:\Crf3\RULE60\10717473.raw

Output Set: N:\CRF4\08052004\J717473.raw

L:26 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:27 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:33 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1) (vi)  
L:68 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1